コーパス研究と中学校英語教科書
－新しい動詞分類基準に向けて－

言語教育体系論講座
岡　田　　毅

Abstract

This paper explores the possibility of integrating corpus research with EFL teaching in Japan. Given that conversational skills are primary objectives in EFL teaching in Japanese junior high schools, the 'language elements' in textbooks should appear in a more systematic way based on quantitative evidence. The official document of the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) entitled 'The Course of Study for Lower Secondary School' (2003) requires that the 'language elements' including words, idioms and grammatical items should be taught in accordance with their 'levels of difficulty', i.e., from easy to more difficult. However, since it is not feasible to define what is 'easy' or 'more difficult' to learn, statistics of frequency will most likely provide the most tangible evidence. We focus on twelve common English verbs proposed in Biber *et al.* (1999) and investigate the frequency of six conjugational forms of each verb in the British National Corpus (BNC). A close examination of the conjugational pattern of each verb in the BNC reveals that, for example, although it is intuitively easy to imagine that the communication verb *say* is one of the commonest ones, it is not its finite base form, but its past tense form *said* that is used most frequently. Such findings will suggest that the order of appearance of verbs and their conjugational forms in Japanese junior high school textbooks should be modified in a way which reflects the actual language use. In the latter part of the paper we will touch on the possibility of establishing a new criterion for categorising English verbs with reference to the part-of-speech (POS) tags assigned to the conjugational forms.

１．序論

本論では、英語主要動詞のBNC（*The British National Corpus*）中での出現パタンをそれらの活用形の視点から詳細に分析し、そこから得られた知見が「英語での実践的コミュニケーション能力を育成するため、実際の言語の使用場面や言語の働きに十分配慮した」教材を取り上げるべき、とする『学習指導要領』（以後『要領』）に準拠した中学校英語検定教科書において、どうのように有効に反映されるべきかという問題を考察する。[1]

日本における英語教科書で用いられている動詞の分析では、従来はlemmatisationを行い、いわゆる「基本形」リストを作成し、学年別の調査をするという手法が大半を占めていた。下の図１は、㈱イー・キャストの「単語レベルチェッカー Ver.6.3」を用いて中学校と高校で用いられている英

語教科書中で、*give*という動詞がどのように扱われているかを調査した結果の画面である（http://www.e-cast.jp/tangochk.htm）。学年別のセルの下の数値が示すのは、この動詞の「認知度」とされている。中学校・高校の各教科書には、それぞれの市場シェア率があり、必須単語については、語数が定められているが、教科書によって取り上げる単語と学年は異なる。従って、生徒の何%が当該の単語を学習しているかも、教科書のシェア率や学年によって異なる。 中学校の全教科書計18冊と高校の教科書計14冊をシェア率と学年に準拠して調査し、当該単語の認知率を学年別に数値化したものが示されている。[2]

図１　「単語レベルチェッカー Ver.6.3」の出力

単語の認知率＆レベル

give

中1	中2	中3	高1	高2	ﾚﾍﾞﾙ
3.9	100	100	100	100	中2

単語の認知率＆レベル

gave

中1	中2	中3	高1	高2	ﾚﾍﾞﾙ
3.9	100	100	100	100	中2

単語の認知率＆レベル

giving

中1	中2	中3	高1	高2	ﾚﾍﾞﾙ
3.9	100	100	100	100	中2

単語の認知率＆レベル

given

中1	中2	中3	高1	高2	ﾚﾍﾞﾙ
3.9	100	100	100	100	中2

単語の認知率＆レベル

gives

中1	中2	中3	高1	高2	ﾚﾍﾞﾙ
3.9	100	100	100	100	中2

注意すべきは、原形・過去形・-*ing*形・過去分詞形・3単現形のいずれを対象語としてもこのシステムが全く同じ数値を返し、全ての形が中学校の第1学年での3.9%の生徒に認知され、第2学年以降は全ての生徒が教科書中でこれらに接していると明示していることである。後続の議論で明らかになるが、これは事実を反映しておらず、この数値は5種類の活用形を無視して得られた結果である。つまり、当該動詞の基本形がどの学年の教科書に現れたかを示しているしか過ぎず、実際の動詞の形およびそれらが用いられた構文情報は考慮に入れられていない。このようなlemma変換後の調査とは対照的に、この研究では英語の主要動詞それぞれの6つの活用形ごとの出現傾向を、教科書別・学年別に分析する。

(1)　BNCにおける一般動詞に与えられる活用を示す品詞標識（Part-of-Speech Tag）

VVB	条件法現在 (命令法を含む)	VVD	過去形
VVG	-*ing* 形 (動名詞/現在分詞 の区別なし)	VVI	原形不定詞
VVN	過去分詞形	VVZ	-*s* 形

教科書における動詞活用形の導入順序としては、例えば現在形のような「基本形」から過去分詞形のような「発展形」へという流れや、平叙文、疑問文、否定文から受動文や完了時制を伴った文、というような、文部科学省のいう「学習段階に応じて平易なものから難しいものへと段階的に指導」

する、という順序が受け入れられているが、この難易度に厳密な尺度が存在しない以上、大規模で包括的なコーパス中の出現頻度情報が、導入順序に対しても重要な意味を持つのではないかという主張を本論では展開する。BNCのような英語を包括的に扱うコーパスの分析結果を、学習教材としての中学校英語教科書の傾向と直接比較することについての問題点などが本論での議論を通して提起できることが望まれる。これは教科書編纂時に参照すべきコーパスはどのようなものが相応しいかという議論にも繋がるからである。

また、英語母語者の第1言語習得過程を反映するような性格のコーパスやデータ（例えばCHILDES（Child Language Data Exchange System）の分析によって得られる成果を、第2言語ないし外国語として英語を習得し、その「実践的コミュニケーション能力」を育成するために用いられる教科書の編纂にどのように組み入れるべきかという議論も今後必要となる。この際に、学習者コーパス（Learner Corpus）から得られる学習者の生産物の実態に関わる知見からのフィードバックも構造的に位置づけられなければならない。この点において、2007年8月に一般公開の開始された日本人中学生・高校生による英作文データであるJEFLL Corpus（http://scn02.corpora.jp/~jefll03/jefll_top.html）は大きな可能性を持つと考えられる。

本論後半では、研究の課程において浮上する動詞に付与されるBNCの品詞標識6種類の妥当性について議論を加え、動詞の分類基準に向けての新たな可能性を示唆する。動詞の「基本形」はCLAWS1タグセットにおけるVBに相当するが、BNCでは形態上同一の動詞がVVBとVVIとに区分されている（Burnard（ed.）2000: 75）。このうちVVIはinfinitive form of lexical verbsを示すが、それは『要領』でいうところの「不定詞」と完全に同義ではなく、⑵に示すような、否定辞や助動詞に先行されたり疑問文における語順倒置の結果生じたtenseless verbsと考えるべきである。

⑵ A05 I like to **think** that he would have accepted that art is work...
A06 No, I don't **think** so.
A06 Siegfried, I don't know what you'll **think**, but I've been recommended...
A06 What do you **think** existing kinds of dram training do for the new actor?

否定文や疑問文などは中学校の早い段階で導入される構文であり、ここで語形の変化を伴わずに用いられる動詞を「不定詞」としては教えない。従って、中学校英語教科書中の動詞活用形分布を調査分析する際に、VVIという品詞タグは適切ではない。しかも、BNCにおける動詞分析においてもVVIは無標のタグとみなすべき根拠が数量的にも存在する。この考え方に従って英語主要動詞の新しい分類についても後に言及する。

２．中学校教科書における「実際の用法」

Römer（2004:156）が指摘するように、EFL（English as a Foreign Language）学習者のコミュニケーション能力を高めるのに役立つ教材の必要性を強調するのは、日本のみならず全世界的に見られる傾向である。文部科学省の示す『要領』の詳細な検討が本論の目的ではないが、平成10年12月告示、15年12月一部改正の『要領』の第9節「外国語」の最初に示された、やや広すぎる「目標」に続く下記の項目中の文言は、コーパス研究と大きく関係するものである。

⑶　教材は、英語での実践的コミュニケーション能力を育成するため、実際の言語の使用場面や言語の働きに十分配慮したものを取り上げるものとする。（第2「各言語の目標及び内容等」（英語）　3「指導計画の作成と内容の取り扱い」⑵）

Biber et al.（1998: 4）が主張するように、コーパス研究は経験的な科学で、自然に発生したテキスト中の用法のパタンを分析するものである。BNCのような大規模コーパスは、英語の現実の姿を網羅的に捉えようとするものであり、その意味でこのようなコーパスの分析から得られる数値的・質的な知見はEFL教科書にとって貴重な資料を提供するものと捉えることができよう。

後述する昭和52年7月告示の『要領』には、各学年別に「1 目標」、「2 内容」、「3 内容の扱い」という項目が設けられており、「2 内容」の中は「⑴ 言語活動」と「⑵ 言語材料」に区分されていた。この「言語材料」の項目において、それぞれの学年で取り扱うべき発音・文の種類・文型・語彙・文法事項などが定められていた。しかし、「ゆとり教育」の考え方の影響から、「言語材料」の学年別配置が平成元年3月告示の『要領』で廃止されることになった。[3]さらに平成10年12月告示、平成15年12月一部改定の現行『要領』では、学年別の記述も廃止されて、3年間を通した「目標」と「内容」の提示となっている。また、現行『要領』で特筆すべきは、「言語の使用場面」と「言語の働き」が、言語活動を行う際に十分に取り上げられるべき要素として掲げられ、具体的に教科書で扱い教えるべき言語材料は、これらに沿って、いわば従属する形で選定されるという点である。従来は学年別に配置されていた「言語材料」は、全体的な「目標」を達成するための言語活動を実現するために適宜用いられる項目という性格を帯びることになった。

このために、検定教科書や教員にはより広い裁量の幅が与えられたことになるが、本研究で扱う平成14年度版の教科書には、依然として学年別進行の言語材料の配置という昭和52年度の『要領』の性格が色濃く残されているのである。

現行の『要領』では、中学校の3年間を通じての目標達成を掲げるとともに、⑷で示すように「言語材料」の配置や提示順序について一定の指針が示されている。

⑷　「言語材料」については、学習段階に応じて平易なものから難しいものへと段階的に指導する

とともに、理解の段階にとどめたり表現の段階まで高めたりするなどして効果的に指導すること。（3　指導計画の作成と内容の取り扱い　(1)(ウ)）

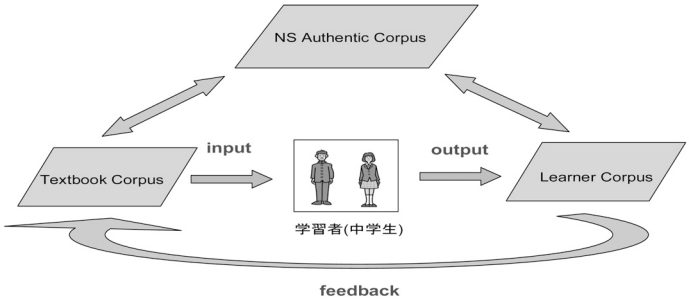
以上を総括すると、学年別に教えるべき言語材料が定められているのではなく、言語の「使用場面」や「働き」を十分反映した内容に応じて、しかも生徒の理解度や学習段階に応じて言語材料が適宜選ばれなければならないということである。

「平易なものから難しいものへ」という段階的な提示は、おおよそあらゆる教育分野に共通する妥当性を持つものであろう。しかし、例えば、昭和52年版の『要領』中の、学年別提示の言語材料・文のように、単文（1年生）→重文・複文（2年生）、言語材料・文法事項のように、現在時制・現在進行形（1年生）→過去形・未来形・過去進行形（2年生）→現在完了形（3年生）という順序が、一見すれば基本形（平易）→展開形（難）という流れに沿っているように思えても、果たして実際にこれが学習者の側からみた適切な段階を踏まえているのかという点に関して、学習者コーパスの分析から得られる知見がほとんど援用できなかった当時の状況を鑑みても疑問を呈さざるを得ないのである。

学習者にとっての難易度に客観的な裏付けがなされない以上、母語話者コーパスから得られる頻度情報が、より「実際の言語の使用場面や言語の働き」に則した言語材料提示の順序に大きな示唆を与え得るのではないかと考えられる。

図2に示すのは、母語話者コーパス、学習者コーパスと並んで、学習者に対する第一義的なインプットとなる教材（教科書）コーパスの相関関係である。日本人の中学生にとって、教科書以外の英語教材に触れる機会は多いといえるが、それでもなお、Römer（2004: 151）の主張するように、インプットとしての大きな比重を教科書が占めているのである。英語教育の大きな目標を達成するべく、言語材料を選出し然るべき順序で教科書中に提示することが求められる所以でもあり、その意味で言語の使用場面や働きに強引に適応させるべく例文などを人工的に作り出すのではなく、実際の母語話者のアウトプットであるコーパスからそれらを収集していくという姿勢が重要になる（Sinclair 1997: 31）。教科書は最早、単に教材のデータとしてだけ扱われるべきではなく、それ自

図２　学習者を中心に据えたコーパスの相関関係



身が独立したコーパスとしての体裁を持ち、固有の研究対象としての位置付けをされるべきなのである。

3．BNC中での主要動詞の活用形パタン

Biber et al.（1999: 373）では、LSWE（Longman Spoken and Written English）Corpus中の英語の主要動詞12個を取り上げ、使用域（register）別の分布を調査している。本節では、BNCで一般（語彙）動詞に付与された品詞タグを基に、これら12個の主要動詞を6種類の活用形の観点から分類する。

⑸　LSWE Corpus中の主要動詞

say，*get*，*go*，*know*，*think*，*see*，*make*，*come*，*take*，*want*，*give*，*mean*

これまでの動詞の分布に関わる研究には、それぞれの動詞のlemmaに着目したものが多く、前述のBiber et al.（1999）もその例外ではない。これに対してOkada（2002）が行ったような、動詞の活用形に着目した詳細な分析により、コーパスの内部構造はもとより、特定使用域に特徴的な動詞の使用パタンを精密に抽出することができる。

表1はBNCにおける主要動詞の活用形別の出現頻度情報を示している。[4]　12個の主要動詞は、それぞれの総出願回数（100万語当たり）順に順位付けされている。それぞれのコラムは各動詞の6つの活用形別のセルに分かれており、上段は総出現回数、下段は総出現回数に対する割合をパーセントで示している。

表１　BNC中における主要動詞の出現頻度情報

count %	say	make	get	go	take	see	know	give	come	think	want	mean	AVG	Total
														STDEV of rate
VVB	237.13	184.41	255.25	219.06	174.52	562.98	568.28	121.18	194.5	477.58	262.66	267.71	293.77	3525.26
	10.32	9.89	14.20	12.29	11.08	36.18	42.06	10.43	16.83	44.76	39.98	45.15	24.43	15.46
VVD	1040.3	236.5	218.6	326.55	281.72	153.26	106.34	163.94	326.77	186.27	120.9	53.21	267.86	3214.36
	45.27	12.68	12.16	18.32	17.89	9.85	7.87	14.11	28.27	17.46	18.40	8.97	17.60	10.34
VVG	147.76	232.49	171.08	495.25	195.02	42.72	27.58	116.83	133.92	43.92	4.98	7.47	134.92	1619.02
	6.43	12.47	9.52	27.79	12.38	2.75	2.04	10.05	11.58	4.12	0.76	1.26	8.43	7.55
VVI	323.23	540.42	569.41	467.23	464.76	431.42	369.33	296.34	240.62	245.25	175.49	92.79	351.36	4216.29
	14.07	28.98	31.69	26.22	29.51	27.73	27.34	25.50	20.81	22.99	26.71	15.65	24.77	5.45
VVN	156.06	503.21	510.02	131.09	336.37	330.24	215.39	351.26	102.33	81.54	13.68	28.18	229.95	2759.37
	6.79	26.98	28.38	7.36	21.36	21.22	15.94	30.23	8.85	7.64	2.08	4.75	15.13	10.11
VVZ	393.69	167.62	72.28	143.01	122.25	35.25	63.58	112.22	157.94	32.58	79.07	143.85	126.95	1523.34
	17.13	8.99	4.02	8.03	7.76	2.27	4.71	9.66	13.66	3.05	12.04	24.26	9.63	6.14
TOTAL	2298	1865	1797	1782	1575	1556	1351	1162	1156	1067	657	593		
STDEV of rate	14.61	8.90	10.96	8.92	7.97	13.94	15.45	8.95	7.04	15.84	15.06	16.18		

各コラム最下段には、当該動詞の総出現回数と各活用形の割合の標準偏差が示されている。最右側のコラムは、活用形ごとの総出現回数と、それぞれの活用形の割合の標準偏差を示す。右から2つ目のコラムは出現回数、割合の平均値を示している。また、網掛けになっているセルは、その活用形が当該動詞の総出現回数に占める割合が25%以上のものである。

各動詞の総出現回数ではなく、特定の動詞の6種類の活用形のうち何が最頻出かを論じることに

より、より精密なテキスト範疇化も可能になる。また、本論でのように、学習者へのインプット・コーパスとしての英語教科書中に、どのような活用形をどのような文型や構文の中で提示するべきかという議論に展開することが可能となる。例えば*say*は高頻度の動詞であると想像することは非英語母語話者にとっても困難なことではないが、*say*の出願回数を大きく引き上げているのは、過去形の*said*であり、過去分詞形の*said*や*-ing*形の*saying*は少ない、という数量的なプロポーションを直観的にいい当てることは難しいのである。一方で、*mean*という動詞が、*'I mean...'*などの慣用的な表現の頻出のせいで、VVB形（定形の動詞原形）で最も突出するということは想像できるが、BNCのSpokenパートでは確かに89.82%と圧倒的にVVB形が多くとも、WrittenパートではVVBはわずかに12.55%で、不定形の*mean*が29.3%、三人称単数現在の*means*が25.91%と多い、というような事実もこのような活用形別の分析をしなければ抽出できないのである。[5]

偏って多用される傾向のある特定の慣用表現の扱いや、SpokenパートとWrittenパートの間での対比は、コーパス研究プロパーの魅力ある領域であり、稿を改めて詳細に検討する。本論では、表1に準拠して、それぞれの動詞を特定の活用形を強く示す項目として性格付けし、中学校教科書での初出学年を7種類の教科書別に横断的に分類する。

４．主要動詞の英語教科書中での出現傾向

本論での分析対象である平成14年度版、7種類の検定教科書、それぞれの教科書名・出版社名・表の中での略称等は以下のとおりである。

(6)

教 科 書 名	出版社	略 称
New Crown English Series	三省堂	N Crown
Columbus 21: English Course	光村図書	Columbus
New Horizon: English Course	東京書籍	N Horizon
One World: English Course	教育出版	One World
Sunshine: English Course	開隆堂	Sunshine
Total English	学校図書	Totl Eng
Total Active. Comm: English for Active Learners	秀文館	Totl Actv

表2は、lemmatiseした主要動詞が初めて用いられた学年を教科書別に示したものである。各動詞の列はそれぞれ3分割されており、第1学年～第3学年を表わしている。網掛けで示したセルは、当該動詞が全く用いられていない学年を教科書ごとに示したもので、例えばBNCにおいて総出現回数最多の*say*は*New Crown*、*New Horizon*、*Total Active. Comm*の1年生の教科書には、いかなる活用形ででも現れていないことを表わしている。7種類の教科書の第1学年用全てに用いられているのは、*go*，*see*，*know*，*come*のみである。とりわけ*give*は全ての教科書の第1学年用では一切用いられていないし、*mean*などは、第2学年の教科書においてでもあまり用いられていない。

本論で焦点を当てるのは、lemmatiseした動詞の出現傾向ではなく、それぞれの活用形であるた

めに、表2の最左列は空白になっている。この列に各動詞が典型的に有する活用傾向を示す品詞標識を入れ、なおかつ当該活用形の初出学年を示したものが表2である。表3の中で、▭マークで濃い目の網掛けとなっているセルは、表2では○であった、つまりlemmaとしては取り上げられていたのに、問題となっている動詞の最多活用形は現れていない教科書と学年を示している。逆にいえば、最多活用形以外の形でその動詞が用いられたということである。例えば、*say*の最多活用形である過去形の*said*が第1学年で現れているのは、⑺のように、*One World*だけということになる。[6]

⑺　"I am very hungry," **said_VVD** Wasika.（*One World*　Year 1, p.86）

表 2　主要動詞（そのlemma）の別初出学年

		N Crown	Columbus	N Horizon	One World	Sunshine	Totl Eng	Totl Actv
	say	×	○	×	○	○	○	×
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	make	○	○	×	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	get	○	○	○	○	○	○	×
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	go	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	take	×	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	see	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	know	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	give	×	×	×	×	×	×	×
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	come	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	think	○	×	×	○	×	×	×
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	want	×	×	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
	mean	○	×	×	×	×	○	○
		○	×	×	○	×	○	○
		○	○	○	○	○	○	○

これは、この年度の教科書が従うべき『要領』には、もはや学年別の文法項目指定がなされていないにもかかわらず、2代前に当たる昭和52年の学年別区分のある『要領』の影響を強く反映していると考えざるを得ない。そこでは、動詞の過去時制は第2学年で教えるべき言語材料として指定されている。*New Crown*、*New Horizon*、*Total Active*では*say*そのものが用いられていないが、*Columbus*、*Sunshine*、*Total English*では*say*は⑻に見られるように、VVD以外の活用形で現れている。

⑻　**Say_VVB** "please", Nick.（*Columbus*　Year 1, p.38）

表 3　各主要動詞にとって典型的な活用形の初出学年

		N Crown	Columbus	N Horizon	One World	Sunshine	Totl Eng	Totl Actv
		×	–	×	○	–	–	×
say	said_VVD	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		–	–	×	–	○	○	○
make	make_VVI	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		–	–	○	○	○	–	×
get	get_VVI	–	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	–	–	–	–	–
go	going_VVG	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		×	○	○	○	–	–	○
take	take_VVI	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	○	○	○	○	○	○
see	see_VVB	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		–	–	○	–	○	–	○
know	know_VVB	–	○	○	○	○	○	○
		○	○	○	○	○	○	○
		×	×	×	×	×	×	×
give	given_VVN	–	–	–	–	–	–	–
		–	–	–	○	–	○	○
		–	–	○	○	–	–	–
come	came_VVD	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		○	×	×	–	×	×	×
think	think_VVB	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		×	×	–	○	○	○	○
want	want_VVB	○	○	○	○	○	○	○
		○	○	○	○	○	○	○
		–	×	×	×	×	○	–
mean	mean_VVB	○	×	×	–	×	○	–
		○	○	○	○	○	○	–

Simon **says_VVZ** ,“Touch your hands.”（*Sunshine*　Year 1, p.41）

How do you **say_VVI** pan in English?（*Total English*　Year 1, p.67）

VVD型の性質を示すもうひとつの動詞項目である*came*もまた、2種類の第1学年用教科書で取り扱われているだけである。

(9)　I **came_VVD** back yesterday.（*New Horizon*　*Year 1, p.97*）

In the evening Maria’s father and brother **came_VVD** home.（*One World*　Year 1, p.82）

VVIは、序論で指摘したように、第1学年で全て取り上げられるべきなのに、表3中の*make*, *get*, *take*の濃・淡網掛けのセルからは、必ずしもそれが正しくはないことが示されている。これは昭和52年の『要領』の中で、「不定詞」は第2学年で初めて教える項目として指定されているためと考えられる。[7]

学年別区分のあった『要領』において、動詞の進行形は現在時制と組み合わせて第1学年で、過去時制と組み合わせて第2学年で、というように早い時期から教えるべきと指定されている。現行の『要領』では学年別記述がないために、時制に関わりなく進行形が早い時期から取り上げられて

いる。これは特に教室内外でのALTとの会話の際に進行形を自然に用いることができるようにという配慮からなされていると考えられる。しかし、Spokenにおける進行形の量的な特徴に関してLeech（2004）は⑽に示すような興味ある主張をしている。

⑽　The Progressive is increasing in frequency in English, especially in the spoken language. But it is worth noting that the Progressive construction is still much less frequent than the non-progressive Simple Present and Simple Past.（Leech 2004: 18）

もしこの主張が正しいとすれば、中学生の早い段階から進行形に、それも現在進行形→過去進行形という順序で触れさせる積極的な必要性はないのではないかと考えられる。

表1をみても、VVG形は最も出現割合が低く、しかもBNCではこの標識の品詞に関する情報量が豊富とはいえず、VVGが単に動詞の*-ing*形を表わすだけで（Burnard（ed.）2000: 75）、進行形を構成する現在分詞と形容詞的なそれ、動名詞とを区別していないのである。

進行形を示すとは限らないVVGであっても、*go*には最多の活用形であり、*be going to*を中心とした慣用的な進行形がこの特徴をもたらしている。ところが、表3で明らかなように、*going*_VVGは第1学年では⑾で示す2種類の教科書にしか現れない。

⑾　They're **going_VVG** to a supermarket.（*New Crown*　Year 1, p.68）
　　Going_VVG home？（*Columus*　Year 1, p.30）

表 4　given_VVNの前後環境（BNC全体）

be	4142	given_VVN 34074	the	4996
was	2965		to	4699
been	2348		a	3305
is	2262		by	2310
are	1467		him	591
were	1372		her	574
being	600		me	547

had	2208	given_VVN 34074	in	1741
have	1701		up	667
has	1324		for	548
ve	443		them	334
s	256		as	324
d	191		on	324

過去分詞形が割合的に最多となる*give*に関しては、このVVNが完了時制中のものか受動態のものかを解明しなければならない。表4は、*given*の直前の環境を調べるだけでも、それが受動態で多用される活用形であることを示している。

表3からは、*give*という動詞そのものが7種類の教科書の第1学年用で一回も取り上げられていないことと、第2学年になってようやく*give*が現れても、それらの全てがVVNではないことが読み取れる。そして第3学年になりようやく3種類の教科書中で*give*_VVNが現れるが、その内訳は⑿のように、頻出であるはずの受動態の文の中ではない。

⑿ You've **given_VVN** us a lot to think about.（*One World* Year 3, p.67）
They don't just need things **given_VVN** by others.（*Total English* Year 3, p.47）
When this trash is burned, dioxin is **given_VVN** off.（*Total Active* Year 3, p.35）

昭和52年の『要領』において、第2学年で受動態、第3学年で現在完了時制を指定していることに鑑みれば、特に第2学年で*given*_VVNを用いた受動文が一度も取り上げられないということも含めて、*given*_VVNの取り扱い方には疑問が残ると判断せざるを得ない。さらに、時制や態の導入順序に加えて、中学校レベルの教科書では*given*という動詞は、その他動詞としての特殊性ゆえに十分には取り上げにくいのではないかと考えられる。（この議論においてのみ最新の平成18年度採用の最新版を分析した結果を用いる。詳細な論述は機会を改めて展開する。）

典型的にはいわゆるSVOO形とSV DO（direct object）*to* IO（indirect object）形をとる*give*にとってどちらを基本形とみなすかに関しては長年にわたる議論がある（Okada 1988）。そのために、特に受動態でこの動詞を用いる際に、いずれの形から派生した構文を掲げるのかが教科書中では問題となる。

図３　平成18年度採用中学校英語教科書中の*give*

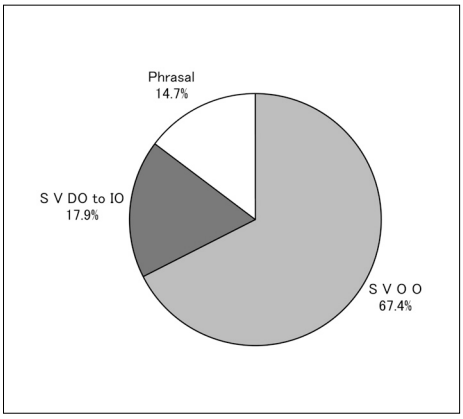
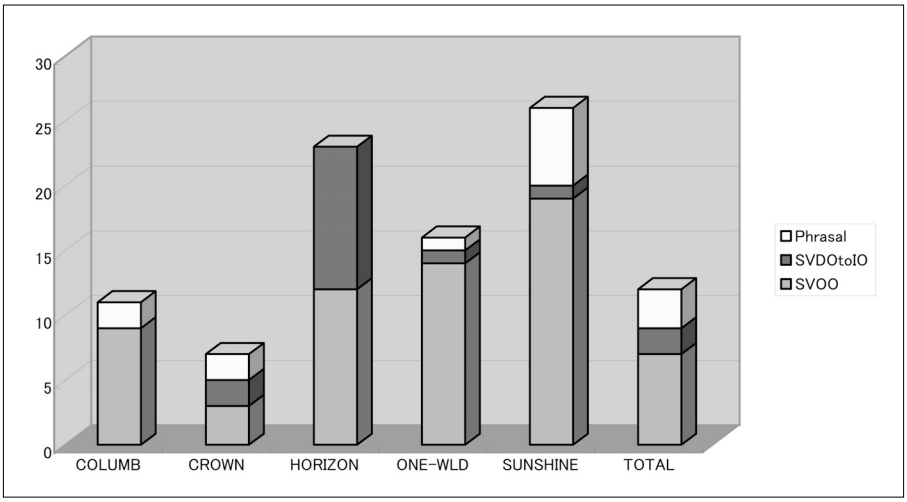


図3は、*given*がSVOO構文、SV DO to IO構文、句動詞のいずれの形で用いられているかの割合を中学校の全教科書を対象に調査した結果であり、図4は6種類の教科書ごとでの出現回数と用法・構文別の割合を示したものである。

５．BNCにおける動詞活用形の分類：その妥当性

先に、『要領』でいうところの「不定詞」は、BNCにおける品詞標識VVIと完全に一致する概念ではないということに言及した。否定文や疑問文などは中学校の最も早い時期から導入される文で

図４　平成18年度採用中学校英語教科書中の*give*（教科書別）



ある。これらの文中で現れる原形の動詞にVVIが与えられるわけだから、表3において第1学年の行にこの活用形で用いられた動詞が含まれるのは当然である。

しかし、このVVIに関して、表1からも興味深い特徴が観察できる。VVIの列の%とその列の右端2つのセル内に注目すると、鮮明なVVDパタンを示す*say*とVVBパタンを示す*mean*を除いて、ほとんどの主要動詞が強いVVI形を示しているということが判明する。割合の平均値も24.77%と高く、標準偏差は5.45と低い。（しかも*say*と*mean*を除いた場合、平均は26.75%、標準偏差は3.15となる。）これらは、動詞の不定形を表わすVVIという標識が、活用形に基づいて動詞を分類しようとする際には有効な指標ではないことを示唆しているのである。

ここで活用形を人間の血液型に喩えてみよう。VVIは血液型分類におけるRH要素と考えることができる。人間の血液型はABO型にこのRH要素型を組み合わせて分類されるのが最も一般的である。日本人の99.4%以上がRH要素を持ち（RH+）、英国人の83.6%がRH+であると言われている。[8] このRH要素に似て、VVIという動詞活用タグは、問題の動詞が時制を伴わないということだけを示し、その意味で、活用形に準拠して英語動詞を分類する際には有効な指標でないと考えられる。

このVVIを無標のタグとみなして、改めてBNCにおける主要動詞の特徴を分析すると、SpokenとWrittenパートで鮮明なコントラストが浮かび上がる。表5が示すように、Spokenにおいては8個ないし9個の動詞がVVBで多く現れる。VVBの全動詞の平均は47.82%と高いのに対して、表6で明らかなようにWrittenでの最高平均割合はVVDである。Spokenで圧倒的にVVBの多い*mean*（88.11%）は、Writtenでは特異なVVZ型を示す。同様にSpokenで強いVVB型を示す*know*は、WrittenではむしろVVN型を示す。

Writtenパートでは一般的にVVDとVVN型が多い。ただし*get*はSpokenで強いVVN型を示す。これは

口語における*'ve got*などの頻出に起因すると推測され、このことは表7からも明らかである。

表 5　VVIを除いた分類（Spoken）

count %	say	make	get	go	take	see	know	give	come	think	want	mean	AVG	Total STDEV of rate
VVB	793.29	257.4	1160.54	956.13	366.86	1310.61	3503.86	310.78	803.64	2623.83	932.82	1911.09	1244.24	14930.85
	19.15	24.57	18.64	20.73	30.94	68.00	89.52	36.56	35.95	72.49	66.77	88.11	47.62	27.36
VVD	1943.77	155	956.8	817.75	259.63	193.39	183.65	138.95	439.96	695.14	252.28	51.83	507.35	6088.15
	46.93	14.79	15.37	17.73	21.90	10.03	4.69	16.34	19.68	19.21	18.06	2.39	17.26	11.09
VVG	537.34	199.58	512.78	1994.73	209.25	57.34	28.24	108.59	463.08	139.73	8.99	6.09	355.48	4265.74
	12.97	19.05	8.23	43.25	17.65	2.98	0.72	12.77	20.72	3.86	0.64	0.28	11.93	12.40
VVN	228.01	281.48	3403.59	398.58	229.08	347.62	83.45	198.71	167.48	100.37	15.76	35.97	457.51	5490.1
	5.50	26.87	54.66	8.64	19.32	18.04	2.13	23.37	7.49	2.77	1.13	1.66	14.30	15.60
VVZ	639.74	154.23	193.29	445.38	120.77	18.28	115.07	93.12	361	60.34	187.3	164.09	212.72	2552.61
	15.44	14.72	3.10	9.66	10.19	0.95	2.94	10.95	16.15	1.67	13.41	7.56	8.90	5.34
TOTAL	4142	1048	6227	4613	1186	1927	3914	850	2235	3619	1397	2169		
STDEV of rate	15.86	5.57	20.30	13.98	7.51	27.66	38.89	10.40	10.32	30.21	27.23	38.17	20.51	

表 6　VVIを除いた分類（Written）

count %	say	make	get	go	take	see	know	give	come	think	want	mean	AVG	Total STDEV of rate
VVB	160.48	174.35	130.48	117.47	148.01	459.94	163.69	95.05	110.54	181.78	170.3	41.22	162.78	1953.31
	9.57	12.80	24.25	13.65	13.46	45.37	28.17	10.96	14.94	41.66	47.97	15.24	23.17	14.25
VVD	915.78	247.73	116.86	258.86	284.76	147.73	99.82	167.38	311.16	116.14	102.79	53.4	235.20	2822.41
	54.63	18.18	21.72	30.08	25.90	14.57	17.18	19.29	42.06	26.62	28.95	19.74	26.58	11.57
VVG	94.07	237.03	123.99	288.59	193.06	40.7	27.49	117.97	88.56	30.72	4.42	7.66	104.52	1254.26
	5.61	17.40	23.04	33.54	17.56	4.01	4.73	13.60	11.97	7.04	1.24	2.83	11.88	9.68
VVN	146.14	533.77	111.22	94.22	351.16	327.85	233.58	372.28	93.35	78.95	13.39	27.11	198.59	2383.02
	8.72	39.18	20.67	10.95	31.94	32.34	40.20	42.91	12.62	18.09	3.77	10.02	22.62	13.96
VVZ	359.78	169.46	55.6	101.34	122.46	37.59	56.48	114.85	136.16	28.76	64.15	141.06	115.64	1387.69
	21.46	12.44	10.33	11.78	11.14	3.71	9.72	13.24	18.41	6.59	18.07	52.16	15.75	11.97
TOTAL	1676	1362	538	860	1099	1014	581	868	740	436	355	270		
STDEV of rate	20.28	11.03	5.57	10.89	8.73	18.35	14.33	13.17	12.59	14.70	19.24	19.05	13.99	

表 7　got_VVNの直前環境のSpoken/Written比較（100万語当たり）

've got	Spoken	1722.44
	Written	38.11
have got	Spoken	133.54
	Written	16.70
's got	Spoken	533.47
	Written	11.67
has got	Spoken	61.01
	Written	8.17

また、VVIを除いた5つの活用形に準拠して、その割合の標準偏差の小さい順に12個の主要動詞をSpoken，Written別に順位付けしたものが図5である。*see*，*think*，*know*，*mean*などの原形を用いた慣用表現がSpokenに多いことが、それらのVVB性を強めていると考えられる。このようなSpokenに特徴的な動詞の活用傾向は、表5、表6の標準偏差自身の平均値の差にも現れている。 Writtenにおいて5種類の活用形に大きな「揺れ」がない反面、Spokenでは特定の動詞がその「揺れ」を大きなものとしている。中学校検定教科書がSpokenへのより大きな比重志向するのならば、学習の早い時期に主要動詞のVVB型が慣用表現の中で用いられている例文を提供するべきである。

図５　標準偏差に基づいた主要動詞の順位付け

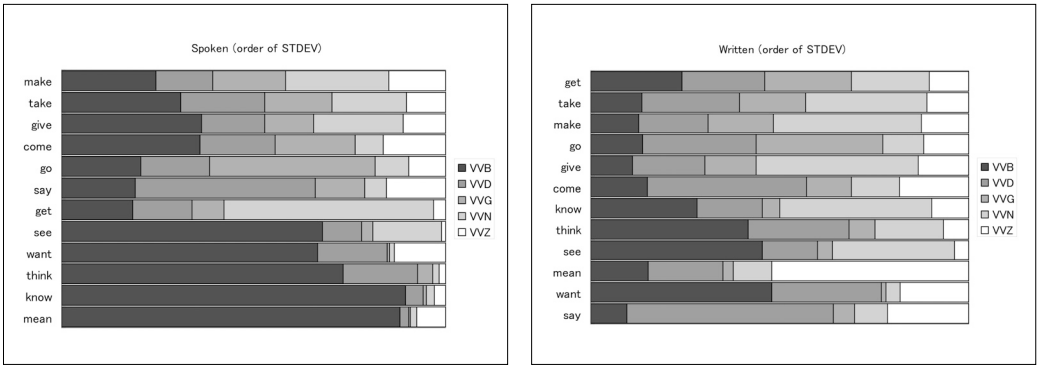


表8は、*I see* や*you know*, *come on*, *I mean*といういような単独で出現し得る慣用的な表現のSpokenとWrittenの対比（100万語当たり）である。*I mean*という表現はSpokenとWrittenで大きなコントラストを示すが、表9が示すように、主語が*I*ではなく、その後に節などを従える表現で用いられる*means*_VVZには鮮明なコントラストは見られない。

表８　慣用表現の中のVVB

	Spoken	Written
I see など	371.99	5.04
you know など	1149.91	10.73
come on など	74.93	2.12
I mean など	376.15	6.84

表９　means_VVZの対比

	Spoken	Written
it means	48.35	15.78
this means	6	32.39
that means	45.35	8.96
which means	17.7	12.93
total	117.4	70.06

６．結論

コーパス研究プロパーの立場でも、単にBNCでは主要動詞のVVI形が多く、VVIは無標のタグに近い、という発見と報告だけでは意味がなく、なぜ動詞の不定形が多いのかに答える必要がある。どのような理由で、つまりいかなる構文の中で動詞の時制や主語の人称や数に関する情報がマスクされてしまい、結果的にtenseless verbとしてのVVI形が多くなるのかを統計的に分析しなければならない。その分析結果は、「VVIを生じさせている構文」を体系的に把握することに繋がり、BNCのような母語話者コーパス中で出現回数の多い動詞のVVI形を、「それが用いられている実際の構文」の中で、実際の教科書中の例文として学習の比較的早い段階に導入する、という教科書編纂手法への貢献となるはずである。動詞の活用形に準拠して文構造を一定の枠組みで規定することができれば、『学習要領』のいう、発達段階に応じた言語材料の有効な提供に資することができると考えられる。

これまで、さまざまな意味で、今後の議論を待たなければならない多くの問題点を指摘してきた。大規模な母語話者コーパスと近年整備の目覚しい学習者コーパスにauthenticityを求めるのと同列に、学習者の第一義的なインプットとなる教科書を精密にコーパス化し、それを分析した上で、他の2つのコーパスとの有機的な関連性を探り、よりよい教科書・教材の開発へと結び付けることが必要であると考える。この際に、教科書が準拠するべき使用域は書き言葉か話し言葉であるべきか、とか、会話・対話の訓練用の素材と、文書を読んだり書いたりするための素材では、その出典となるソースにどのような区分を設定するべきかという問題もある。

加えて、平成18年度4月から採用された新検定教科書でのように、全生徒が必ず学ばなければならない部分と、発展的な情報として従来の『要領』が指定する範囲を大きく逸脱した部分とが織り交ぜられているような場合、教科書コーパス構築にはよりきめの細かい属性の設定が要求される。現在、最新版6種類の全教科書のコーパス化を推進中である。これが整備されれば、どの教科書の何年生のどの課の、しかもどの部分で当該の表現や語彙が用いられているかという情報を精密に抽出し、比較することが可能となる。[9]

1　本論で分析の主な対象とするのは、平成14年度採用の合計7種類の検定教科書である。
2　この最新バージョンは平成18年度採用の検定教科書をデータとしている。
3　昭和52年の改訂版に「コミュニケーション」という文言は1度も現れない。平成元年版で「コミュニケーション」が初めて、1度だけ「第1　目標」の冒頭で用いられた。このために平成元年は「コミュニケーション元年」と呼ばれることがあった。平成10年12月告示の現行『要領』では、「コミュニケーション」は10回も用いられ、「第1　目標」では、「聞くことや話すことなどの実践的コミュニケーション能力」という新しい表現が使われるようになっている。
4　BNC全体から、サブコーパスとしてのImaginativeを除外してある。これは、後の議論で動詞活用形の視点からBNCのSpokenとWrittenパートとの対照に言及するためである。
5　当然のことながらreadingパートを含む教科書が参照すべきはWrittenを無視したSpokenコーパスだけでよいのかというような、話し言葉・書き言葉の対立を反映したコーパスの教科書編纂への援用という重要な問題がここで浮上する。
6　以後の例文では、問題となっている動詞の活用形標識をアンダースコア後に示してある。
7　New Crownの2年生で *get*_VVIが現れない以外は、これら3種類の動詞はVVI形で全ての教科書の2年生で用いられている。
8　常に例外は存在し、バスク地方ではわずか67%の人々がRH+である。
9　平成18年度版教科書の全頁を画像化し、XML形式のコーパスにリンクさせることにより、テキストコーパスには馴染みにくいような情報（例えばイラストに関する記述や写真に映し出されている英語の情報など）を立体的に記述しようと試みている。

参考文献

Aston, G., S. Bernardini and D. Stewart（eds.）（2004）*Corpora and Language Learners.* Amsterdam: John Benjamins.

Biber, D., S. Conrad and R. Reppen（1988）*Corpus Linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press.

Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan（1999）*Longman Grammar of Spoken and Written English.* Harlow: Pearson Education.

Biber, D. and R. Reppen（2002）“What Does Frequency Have to do with Grammar Teaching?” *Studies in Second Language Acquisition* 24, 2: 199-208.

Burnard, Lou.（ed）（2000）*Reference Guide for the British National Corpus*（*World Edition*）.（provided as http://www.natcorp.ox.ac.uk/docs/userManual/urg.pdf）

Hasegawa, S. and K. Chujo（2004）“Vocabulary Size and Efficacy within Three Serial JSH English Textbook: Vocabularies Created in Accordance with Revised ‘Course of Study’ Guidelines.” *Language Education and Technology* 41:141-153.

Leech, G.（2004）*Meaning and the English Verb.* 3rd Edition. Harlow: Pearson Education.

Leech, G., M. Deuchar and R. Hoogenraad（2006）*English Grammar for Today: A New Introduction.* 2nd Edition. Hampshire: Palgrave Macmillan.

Okada, T.（1988）“An Analysis of Double Object Construction in English: Toward a Prototypical Approach to Syntax,” *Collected Treatises Hamamatsu Tanki Daigaku: vol.36,* pp.45-103.

-----（2002）“Conjugational Patterns and Text Categories in the LOB Corpus.” In Saito, *et al.*（eds.）, *English Corpus Studies in Japan.* Amsterdam: Rodopi, pp.43-62.

Römer, U.（2004）“Comparing Real and Ideal Language Learner Input: The Use of an EFL Textbook Corpus in Corpus Linguistics and Language Teaching,” In G. Aston *et al.*（eds.）, *Corpora and Language Learners,* Amsterdam: John Benjamins, pp.151-168.

Saito, T., J. Nakamura and S. Yamazaki（eds.）（2002）*English Corpus Linguistics in Japan.* Amsterdam: Rodopi.

Sinclair, J.（1997）“Corpus Evidence in Language Description.” In A. Wichmann *et al.*（eds.）, *Teaching and Language Corpora.* Harlow: Pearson Education, pp.27-39.

Wichmann, A., S. Fligelstone, T. McEnery and G. Knowles（eds.）（1997）*Teaching and Language Corpora.* Harlow: Pearson Education.

Widdowson, H.G.（1998）“Context, Community, and Authentic Language.” *TESOL Quarterly* 32, 4: 705-716.

（2007年9月25日）